The acquisition order of coherence relations:
On cognitive complexity in discourse

Wilbert Spooren [a,*], Ted Sanders [b]

[a] *Department of Language and Communication, VU University Amsterdam, De Boelelaan 1105,*
NL-1081 HV Amsterdam, The Netherlands
[b] *Utrecht institute of Linguistics – UiL OTS, Trans 10, NL-3512 JK Utrecht, The Netherlands*

Abstract

This article presents an analysis of the acquisition order of coherence relations between discourse segments. The basis is a cognitive theory of coherence relations (Sanders et al., 1992) that makes predictions about the order in which the relations and their linguistic expressions are acquired, because they show an increasing cognitive complexity. The child language literature lends support to two distinctions in the theory, Basic Operation (causal versus additive) and Polarity (positive versus negative). In two studies, additional data were collected to test the validity of two other distinctions, Source of Coherence and Order of the Coherence Relation.

In the first study, children described a picture or conversed freely with the investigator. Both distinctions turn out to be necessary to account for the acquisition patterns. In the second study, the children's proficiency in dealing with negative causal relations was investigated. The two studies use different research designs. The first is a study of relatively naturalistic, only partially structured elicitation of extended stretches of speech produced by children, the other is an experiment on the understanding and production of coherence relations in short sequences of statements relying on nonsense words that lack a conventional semantic content. The two procedures tap very different kinds of communicative skills and linguistic as well as conceptual knowledge. The combination of these two studies allows us to draw valid conclusions about the acquisition of the various coherence relations. The data support the claim that cognitively complex coherence relations show up later than cognitively simple relations.

* Corresponding author at: Department of Language and Communication, VU University Amsterdam, De Boelelaan 1105, NL-1081 HV Amsterdam, The Netherlands. Tel.: +31 20 598 6572; fax: +31 20 598 6500.
 E-mail address: w.spooren@let.vu.nl (W. Spooren).

1. Introduction

Coherence is an essential trait of well-formed discourse. Between the segments of a coherent text, relations like *Contrast* and *Cause–Consequence* must be inferred. These relations play a central role in processes of text comprehension, memory, and reproduction. Consider fragment (1) as an example.[1]

(1)
(a)   De juf is heel aardig
       'The teacher [female] is very nice'
(b)   want bijna altijd zegt ze: ''Goed zo, Deborah''
       'because she nearly always says: ''Well done, Deborah''
(c)   en dat is altijd lief
       'and that's always sweet'
(d)   een meester is heel anders
       'a teacher [male] is very different'
(e)   want die ben ik niet gewend
       'because I'm not used to him/them'

                                                   (Deborah 7;3)

The example shows several things. For instance, the units in this fragment are linked through coherence relations: between clauses (b) and (c) an *Evaluation* relation holds; between (a) and [(b, c)] and between (d) and (e) a *Claim–Argument* relation obtains; the relation between [(a–c)] and [(d, e)] is that of *Contrast*. The example also shows that coherence relations need not be marked explicitly: The *Contrast* relation connecting [(d, e)] to its context has to be inferred by the reader, without the help of explicit linguistic cues like connectives (for example *and*, *but*, and *so*).

In our analysis of this example we use coherence relations and connectives to characterize the discourse connectedness of the fragment. Traditionally, linguists have approached this problem by looking at overt linguistic elements and structures. In their seminal *Cohesion in English*, Halliday and Hasan (1976) describe text connectedness in terms of explicit clues that make a text a text, such as reference, substitution, ellipsis and conjunction. Cohesion occurs ''when the interpretation of some element in the discourse is dependent on that of another'' (Halliday and Hasan, 1976:4).

Over the last decennia, the dominant view has come to be that the connectedness of discourse is a characteristic of the mental representation of the discourse rather than of the discourse itself (see Sanders and Pander Maat, 2006, for a recent overview). According to most cognitive scientists, linguists and psycholinguists, understanding discourse means that readers construct a mental representation of the information in the text (Graesser et al., 1997; Gernsbacher and Givón, 1995; Garnham and Oakhill, 1992). The term coherence is generally used to refer to this type of conceptual connectedness. Discourse coheres in several ways. Generally speaking, at least two types can be identified:

---

[1] Unless indicated otherwise, our examples come from a corpus of child language data collected for Study 1. Notational conventions: utterances were divided into units; a unit is a main clause, an adverbial clause or an appositive relative clause. We have discarded pauses etc. as these seem irrelevant to our analyses. After each fragment, we give the name of the informant and his/her age in years;months.

1. Referential coherence: smaller linguistic units (often nominal groups) may relate to the same mental referent;
2. Relational coherence: A discourse representation shows coherence, because the representation of the discourse segments (most often conceived of as clauses) are connected by coherence relations like CAUSE–CONSEQUENCE

Although there is a principled difference between the cohesion and the coherence approach to discourse, it is important to realize that coherence phenomena may be of a cognitive nature, but that their reconstruction is often based on linguistic signals in the text itself. Both coherence phenomena under consideration – referential and relational coherence – have clear linguistic indicators that can be taken as processing instructions. For referential coherence these are devices such as pronouns and demonstratives, and for relational coherence these are connectives and (other) lexical markers of relations, such as cue phrases and signaling phrases.

In this article we will focus on *relational coherence*. Coherence relations (for the term, see Hobbs, 1979; for their analysis in a formal semantic context, see Asher and Lascarides, 1998) are commonly regarded as ''the cornerstone of comprehension'' (Graesser et al., 2003:82). Examples are relations such as CAUSE–CONSEQUENCE, LIST, and PROBLEM–SOLUTION. These relations hold locally, that is, between adjacent clauses, but also between larger discourse units, from groups of clauses up to whole sections, paragraphs and even chapters (Sanders, 1997; see *Rhetorical Structure Theory*: Mann and Thompson, 1988; Taboada and Mann, 2006). Our view on discourse coherence difference, then, somewhat from Hickmann's (2003), who characterizes cohesion as ''typically defined on the local level of adjacent clauses'', whereas coherence is to be found ''on the more global level of plot organization'' (p. 93).

There have been many studies on the acquisition of connected discourse over the past four decades. Bloom et al. (1980) summarizes a number of findings in the following ordering of the acquisition of relations:

Additive < Temporal < Causal < Adversative

Connectives display the following pattern:

and < and then < because < so < but

Of these, *and* is first used for additive relations, and later also for temporal, causal and adversative relations. The other connectives are used more specifically: *and then* for temporal relations, *because* and *so* for causal relations, and *but* for adversative relations.[2]

In this article we present an explanation for this order, in terms of a theory of coherence relations that is based on the cognitive complexity of general conceptual relations that underlie coherence relations. The theory has been formulated as a classification of coherence relations (Sanders et al., 1992, 1993). In our classification, coherence relations are treated as cognitive entities, as elements of a cognitive construct, the discourse representation. Below, we present an overview of the four parameters underlying the classification (see Appendix A for the complete classification) and some of the findings in the literature that relate to the distinctions in the classification. Then we will report two empirical studies that were carried

---

[2] We disregard the 'place holding' function of *and* and confine ourselves to clause-combining functions of the connectives discussed (Schiffrin, 1986).

out in order to investigate the predictions about the order of acquisition following from the classification.

1.1. An overview of the classification and how it fits acquisition data

For our classification we use four basic relational concepts. They are described in the following sections. For an extensive description, see Sanders et al. (1992, 1993). We also provide an overview of relevant findings in acquisition research.

Basic Operation: additive versus causal relations

The Basic Operation refers to the intuition that discourse segments can either be connected strongly (in causal relations) or weakly (in additive relations). In additive relations there is a relation of conjunction between the two segments (2). In causal relations the Basic Operation is implicational as well as conjunctional (3).

(2)   Additive relation (*List*)
        Zij waren aan het vissen
        en hij was aan het zwemmen.
        'They were fishing
        and he was swimming'

                                                                              (Hanneke 7;8)

(3)   Causal relation (*Result*)
        Daar leunt een dikke man tegen een paal
        en daardoor vallen de blokken van de paal om.
        'There is a fat man leaning against a post.
        and because of that the building blocks of the post fall down'

                                                                              (Daan 11;11)

The results from both naturalistic (e.g., Bloom et al., 1980; Eisenberg, 1980) and experimental studies (Piaget, 1924/1969; Katz and Brent, 1968) of first language acquisition show that additive relations are acquired before causal relations, and that additive connectives are acquired before causal connectives.

Polarity: positive versus negative relations

In a positive relation, the segments are linked directly. In a negative relation the link involves a negation of one of the segments. For example, in (5) the link between the segments relies on the implication relation ''if mothers like something, their children like it as well'', which is negated in the second segment.

(4)   Positive relation (*Explanation*)
        Soms lees ik wel wat
        omdat ik wil weten wat er in het boek staat.
        'Sometimes I read something,
        because I want to know what is in the book'

                                                                              (Bram 7;2)

(5)   Negative relation (*Concession*)
        En mama vindt dat leuk.
        Maar ik vind dat helemaal niet leuk.
        'And mummy likes that.
        but I don't like that at all'

(Deborah 7;3)

The observational studies of Bates (1976), Bloom et al. (1980) and Eisenberg (1980) as well as the experimental studies of, for example, Wing and Scholnick (1981) clearly show that positive relations are acquired before negative relations.

Source of Coherence: content, epistemic and speech act relations[3]

In a content relation, the link involves a real world connection: The relation involves the propositional or ideational content of the related segments (example (6)); in an epistemic relation the connection involves the level of reasoning and concluding (example (7)). In a speech act relation, the connection is at the level of illocutions (example (8)).

(6)   Content relation (*Consequence–Cause*)
        Hier is een boot gezonken
        omdat ze tegen een boei zijn gevaren.
        'Here a ship has sunk
        because they sailed against a buoy.'

(Hanneke 12;0)

(7)   Epistemic relation (*Evidence*)
        Dat zijn aapmensen of mensen van vroeger.
        'Those are apemen or ancient people'
        Want wij zijn in het prehistorisch huis geweest
        'Because we have been in the prehistoric house'
        En daar hebben ze ook van die vellen.
        'And there they also have those hides'

(Bram 7;2)

(8)   Speech act relation
        Nieuws vind ik niet leuk.
        'News I don't find nice' [=I don't particluarly like the news [on television]]
        Want daar vind ik niks aan.
        'Because I don't like it at all'

(Jordy 6;5)

There are two reasons why conclusions about the relevance of Source of Coherence for acquisition data are not straightforward. First, most acquisition studies in the literature use different categories, which makes it hard to interpret the results directly in terms of Source of

---

[3] In our original proposal we distinguish between semantic and pragmatic relations, the latter subsuming both epistemic and speech act relations. For reasons of terminological uniformity we adopt in this article Sweetser's (1990) terminology. See Sanders and Spooren (2001) for an argument that our bifurcation maps on Sweetser's trifurcation in this way.

Coherence. Second, for the categories that can be related to the content-epistemic-speech act distinction, some studies (Peterson, 1986; Bloom and Capatides, 1987; Bates, 1976) seem to show that content relations are acquired earlier, whereas other studies (Kyratzis et al., 1990; Piaget, 1923/1971; Corrigan, 1975) suggest that non-content relations (especially speech act relations) are acquired earlier.

Order of the segments: basic order or non-basic order

Relations are presented in basic order if the information in the first segment corresponds with the antecedent of the Basic Operation; relations are in non-basic order if the information in the first segment corresponds with the consequent of the Basic Operation. For instance in (9) the Basic Operation is a causal one ''someone playing beautifully can cause people to cry''. The first segment relates to the first element in this causal link, the second segment relates to the second element. In example (10) the same Basic Operation is used, but now the links to the causal connection are reversed: The first segment relates to the second part of the causal link (the consequent), the second segment relates to the antecedent.

(9)   Basic order relation (*Cause–Consequence*)
        Een man speelt viool
        'A man plays the violin'
        en iedereen begint te huilen [. . .].
        'and everybody starts to cry [. . .].'

                                                                              (Bas 12;0)

(10)   Non-basic order relation (*Consequence–Cause*)
        Die mensen huilen allemaal
        'Those people are all crying'
        omdat die vioolspeler zo mooi speelt.
        'because that violin player plays so beautifully.'

                                                                              (Jory 7;4)

Order of the Segments does not distinguish between different types of additive relations, since additive relations are logically (although not pragmatically) symmetric.

Prior research shows that in specific, experimental, situations, children make more errors identifying non-basic order relations than basic order relations (Bebout et al., 1980). This suggests that basic order relations are acquired earlier. However, in more natural conditions (e.g., observational studies like Hood and Bloom, 1979; French and Nelson, 1985; McCabe and Peterson, 1985) such a difference is not found. This suggests that there is a confounding with research methodology. Source of Coherence seems to be another relevant factor accounting for this difference between observational and experimental studies: experimental studies have focused on content relations, whereas non-content relations (especially speech act relations) are more frequent in observational studies. A final relevant factor for the difficulties children have with non-basic order relations in experimental studies may be the cognitive load on the participants: Peterson and McCabe (1987) have shown that the children performed better after having carried out the task a number of times or when they could use a picture as a memory aid. In short, this survey shows that research results with respect to order of segments are inconclusive.

1.2. How a theory of coherence relations can account for these data

We want to suggest that the differences in emergence reported in the literature can be accounted for in a systematic way. In our analysis, some relations are considered more complex than others as they involve more complex concepts. For instance, *Concession* relations are more complex than *List* relations because they involve negation on top of additivity, and *Cause–Consequence* relations are more complex than *List*-relations because they involve causality in addition to additivity. From the cognitive claims attached to the analysis, it follows that such complexity patterns show up in the order in which coherence relations are acquired. That is, one might expect *List* relations to be acquired before *Contrast* relations and before *Cause–Consequence* relations.

Similarly, epistemic relations are complex, as they require abstract reasoning, usually exploiting relations that exist between real world events. The conclusion in (11) that this man is a cowboy exploits the real world relationship that cowboys use lassos to catch horses.[4]

(11)　　Dat is een cowboy
　　　　'That's a cowboy,'
　　　　die wil paarden vangen
　　　　'who wants to catch horses'
　　　　want hij heeft een lasso.
　　　　'because he has a lasso'

　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　(Deborah 7;3)

As to the complexity of content versus speech act relations, Kyratzis et al. (1990) argue that speech act relations are least complex.

With respect to the fourth distinction, between basic and non-basic relations, basic order relations are less complex than non-basic order relations, at least for the content domain: events connected by a Cause–Consequence relation iconically reflect the order of the events in the real world.

These considerations lead to two types of predictions:

Hypothesis 1. Weak prediction:
Developmental trends cannot be described without the distinctions made by the classification.

Hypothesis 2. Strong prediction:
Relations that are analyzed as simple are acquired before relations that are analyzed as complex.

In order to test these hypotheses we conducted two studies, in which the relevance of the basic concepts was central. The first is a corpus study of Dutch, in which we compared coherence relations produced by children more or less spontaneously in a partially structured setting. The corpus study was inconclusive with respect to the acquisition of the most complex relations, namely the negative causal ones. Therefore, we conducted an experiment to test children's capacity to use these. In sum, we have used two different research designs. The first is a study of relatively naturalistic, only partially structured elicitation of extended stretches of speech

---

[4] Analyses of epistemic relations in the mental space framework, like Dancygier and Sweetser (2000) and Verhagen (2000), inevitably portray such relations as more complex than their non-epistemic counterparts.

produced by children, the other is an experiment on the understanding and production of coherence relations in short sequences of statements relying on nonsense words that lack a conventional semantic content. The two procedures tap very different kinds of communicative skills and linguistic as well as conceptual knowledge. The combination of these two studies allows us to draw valid conclusions about the acquisition of the various coherence relations.

2. Study 1: children's use of coherence relations while describing and conversing

Our first study is a corpus analytic study under experimental control, using Dutch children as informants. The children were given two tasks, both aiming at the elicitation of complex utterances: a description task (biasing for content relations) and a conversation task (biasing for epistemic and speech act relations). The research was carried out with elementary school children from different age groups: grade 1 (age 6–7) and grade 6 (age 11–12). This cross-sectional design allows inferences on the developmental patterns in the children.

The general research question is to what extent the classification explains for differential patterns in the acquisition data, in accordance with the weak and the strong prediction. The weak prediction is that differences in use of coherence relations between the two age groups can be described in terms of the primitives of the classification. According to the strong prediction, coherence relations that are classified as cognitively simply are acquired before complex coherence relations. More specifically this leads to the following four research questions.

1. Is there any evidence for the strong prediction based on the classification of relations, namely that complex relations are acquired later than simple relations?

The observations of Bates (1976), Bloom et al. (1980), and Eisenberg (1980) and the experimental data of Piaget (1924/1969), Katz and Brent (1968), and Scott (1984) support the distinctions based on the concepts Basic Operation and Polarity. They also suggest that less complex relations are acquired earlier than complex relations. Therefore, we expected younger children to use less negative and causal relations than older children.

2. Is there any reason to assume a difference in degree of complexity between content and other relations?

Different researchers arrive at different conclusions concerning the degree of difficulty of content versus other relations. Diachronically, epistemic and speech act relations are derived from content relations (cf. Sweetser, 1990; Tabor and Traugott, 1998). On that basis it is tempting to assume that non-content relations are more complex than content relations (following Bates, 1976). But ontogeny need not replicate phylogeny (Traugott, 1997). The most sophisticated picture is presented by Kyratzis et al. (1990), who claim that speech act relations are acquired before content relations, which in turn are acquired before epistemic relations. However, these findings are based on data collected in a context that is heavily biased toward non-content (more specifically, speech act) relations. An unanswered question is whether the same goes in a context biased for content relations. To answer this question we will compare the amount of content and non-content (epistemic and speech act) relations produced by the two age groups.

3. Is the Source of Coherence determined by the context (text type) in which the utterance was produced?

The seemingly contradictory results with respect to Source of Coherence may be attributable to a difference in context in which the data were collected. We have argued elsewhere (Sanders, 1997) that text type and the type of coherence relations found in a discourse correlate strongly. Informative and descriptive texts show large numbers of content relations, whereas epistemic and speech act relations dominate argumentative texts. This being so, the type of context in which data are collected seems a major factor in the type of relations produced by the children. We expected to find more content relations in the descriptive task, and more non-content (epistemic and speech act) relations in the conversation task.

4.   Do we need a separate basic concept Order of the Segments, apart from Iconicity?

This question pertains to the discussion whether the Iconicity principle suffices to account for the ordering patterns in the data or whether we need an additional principle, Order of the Segments. The Iconicity principle (or Order of Mention principle, Clark and Clark, 1977:358) states that iconic ordering is more easy than non-iconic ordering, because it reflects the order of the events 'in the real world'. As Iconicity applies to events only, it is restricted to content relations. The crucial question is whether non-content (epistemic and speech act) relations also show preferences for specific orderings. If so, we need an additional principle Order of Segments, apart from Iconicity.

2.1. Method

2.1.1. Materials
Each child was given two tasks: a description task and a conversation task. In the description task (which biased for content relations) children had to describe a number of 'Where's Waldo?' (Handford, 1991) pictures, that show large numbers of causally related events, such as a man walking under a ladder, whistling and unaware that on the ladder a window cleaner has just dropped a bucket of water. In the conversation task (bias for epistemic and speech act relations) the children had to formulate an opinion concerning a number of controversial topics, such as vetoed TV programs, wearing brand clothing at school, the amount of the children's pocket money, etc.

2.1.2. Participants
The investigation was carried out with six children from grade 1 and six children from grade 6 of an elementary school in a medium-sized town in the south of the Netherlands. The choice of these age groups was inspired by choices made in related research: Younger children do not produce wide varieties of coherence relations, and as to the upper boundary of grade 6, Piaget (1924/1969) claims that at age 11–12, children begin to use even the most complex relations. Indeed, Evers-Vermeul (2005) finds hardly any use of these negative causal relations in her study of young children (roughly 1;6–5 years of age), and if they show up, they appear after positive causal relations emerge.

The two age groups in our study consisted of equal numbers of boys and girls. The reading level of the children varied equally in each age group. The mean age of the children in the younger age group was 7;2 (min. 6;7, max. 7;8). The mean age of the children in the older age group was 12;1 (min. 11;11, max. 12;7). All children had Dutch as their first language.

2.1.3. Procedure
Each child was interviewed individually in a separate room of the elementary school they attended. Each child was given the two tasks, each task lasting circa 10 min. In the description

task, the children were given a picture from a *Where's Waldo* book, along with the following instruction: ''In this picture all sorts of things are happening. Can you tell me what is happening and how that has come about?'' In order to prevent a bias from this formulation for specific orderings of descriptions, the formulation of the instruction was varied with ''Tell me how this has come about, what is happening here?'' In the conversation task, the children were stimulated to give their opinion on a number of topics and to motivate their opinion. The responses of the researcher were restricted to presenting the instruction, and to back-channel behavior.

2.1.4. Data analysis

The children's utterances were recorded on a cassette recorder and transcribed. The unit of transcription and analysis was the clause: main sentences, adverbial subordinative sentences and appositive relative clauses were taken as minimal units. Only uninterrupted utterances were taken into account.

Subsequently, the transcripts were analyzed by two judges, on the basis of the set of coherence relations specified in Mann and Thompson's (1988) Rhetorical Structure Theory (RST). This model of analysis was chosen because of its descriptive width. Of each utterance it was decided between which clauses an RST-relation existed. Then this relation was analyzed in terms of the four basic concepts of Sanders et al. (1992). Appendix B gives an example of a protocol fragment. Examples (2–11) above are taken from the corpus and illustrate the categorization. Appendix C shows further criteria for distinguishing between content, epistemic and speech act relations.

Following Bloom et al. (1980) we have used two strategies to guarantee the reliability of the analysis. The coders have made use of all the contextual information available to interpret the relations between the utterances. Furthermore we have the principle of cumulative semantic complexity: the relations have been interpreted as specifically as possible. In terms of the classification this means that additive or temporal relations which can have a causal interpretation, are coded as causal relations.

Note that by following this procedure we are studying the occurrence of coherence relations as well as their linguistic marking. It can be asked whether we study emergence or acquisition. By looking at the use of the relations, it follows that our prime focus is emergence. However, we also investigate acquisition because we study whether coherence relations are correctly used, that is, we pay attention to the (mis)match between connective and coherence relation. Occasionally the children's use of a connective is deviant. An example is (12).

Example (12)

[Wat doe je het liefste als je uit school bent en waarom?]
[What do you like to do most when you are out of school and why?]

1   Nou ik zit op voetbal
       Well I play football

2a   En dat vind ik leuk
       And that I like

2b   omdat ik iedere dag oefen voor voetbal
       because I practice every day for football

3   En dan word ik steeds beter. [cause]
       and then I am getting better and better

4   Dan spelen we echte wedstrijden [list]
       Then we play real games

5   En dan zit het hele stadion vol [list]
       And then the stadium is crowded
6   En dan verliezen we soms. [list]
       And then we loose sometimes
7   Maar het meeste winnen we.
       But mostly we win

[3–1; Jordy; 6;7]

In this fragment Jordy uses a strategy commonly found among young children: many utterances are connected by the temporal marker *en dan* ('and then'), suggesting a temporal sequence of events. However, sometimes a differently relation seems intended. For example, in line 3 *en dan* is used to signal a causal relationship (he probably intends to express a complex relationship that another reason why he likes football is that because of practicing every day he gets better and better at it). The example has been analyzed as a causal relation, signaled by a temporal connective (a so-called implicit use of connectives, see Spooren, 1997). This is an indication that Jordy has not acquired the intricate possibilities of using *en dan* (literally a temporal connective) to express causal relations. Still, he does know how to use the intended causal relationship.

In total, the children produced 1355 utterances (112.9 utterances per child), 756 in the description task (63.0 per child) and 599 in the conversation task (49.9 per child). 1097 coherence relations were analyzed (526 in the description task, 571 in the conversation task). The children in grade 1 produced 566 coherence relations (mean of 94.3 relations per child), the children in grade 6 produced 531 relations (mean of 88.5 relations per child).

Per child, the percentage of categories mentioned in the hypotheses and research questions was calculated (e.g., the percentage of content relations in basic order). These percentages were analyzed using:

a. the Mann–Whitney rank order sum test for differences between two age groups (for instance, the difference in percentage of content relations between grade 1 and grade 6);
b. the binomial test for evaluating one-sample distributions (for instance, the observed percentage content relations compared to the expected percentage if content and non-content relations were to occur equally frequently).

In addition, for the analysis of frequencies χ^2-tests were used.

2.2. Results

2.2.1. Reliability of the analyses
The analyses were carried out by two analysts independent from each other, after careful instruction and on the basis of a coding scheme (see Appendix C for an example of the coding scheme). Cohen's kappa coefficient of agreement between two analysts was calculated separately for each child's data. Agreement was lowest for Order of the segments (with *K* varying from .40 to .83 in grade 1, and from .78 to 1.00 in grade 6). For Basic Operation *K* varied from .66 to .91 for grade 1 and from .88 to 1.00 for grade 6. For Source of Coherence *K* varied from .66 to .88 for grade 1, and from .90 to .98 for grade 6. Finally, for Polarity *K* varied form .54 to 1.00 for grade 1 and from .87 to 1.00 for grade 6.

Table 1
Coherence relations in our corpus classified as content, epistemic or speech act (both by grade and by task) (row percentage in parentheses) (Study 1)

Type of relation	Content	Epistemic	Speech Act
Grade			
1	375 (67.4)	21 (3.8)	160 (28.8)
6	279 (53.9)	37 (7.1)	202 (39.0)
Task			
Description	487 (94.2)	26 (5.0)	4 (.8)
Conversation	167 (30.0)	32 (5.7)	358 (64.3)

Note. In 23 cases the epistemic/speech act character of the relation was unclear.

2.2.2. Tests of the hypotheses

COMPLEX RELATIONS. It was expected that causal, negative and negative causal relations, which – in terms of the classification – are more complex than their additive and/or positive counterparts – are more frequently used by older children. This turned out not to be the case: Causal relations: grade 1: 56%, grade 6: 50% ($z = 1.28$, $p = .28$); negative relations: grade 1: 6%, grade 6: 10% ($z = 1.44$, $p = .16$); negative causal relations: grade 1: 2%, grade 6: 2%; $z = .16$, $p = .88$).

SOURCE OF COHERENCE AND COMPLEXITY. The data are summarized in Table 1. Summed over both tasks, the children produced more content relations than non-content relations (content relations: 59%; binomial test: $p < .05$). Moreover, younger children produced more content relations than older children (grade 1: 66% content relations; grade 6: 52% content relations; $z = 2.56$, $p < .01$).

If we compare epistemic or speech act relations separately, the two age groups show no difference: grade 1 produced 181 non-content relations, 21 of which (11.6%) were epistemic. Grade 6 produced 239 non-content relations, 37 of which (15.5%) were epistemic ($z = .48$, $p = .63$). Therefore, there is no indication that younger children have more difficulty with epistemic relations than older children.

ROLE OF CONTEXT. As expected, it was found that the task determines the type of relations produced by the children: in the descriptive task more content relations (92.6%), in the conversational task more non-content (epistemic and speech act) relations (72.3%) (in a binomial test, both percentages differ from chance, $p < .001$).

As can be seen in Table 1, the predominance of non-content relations in the conversational task is mainly due to the number of speech act relations: in the description task there are 30 non-content relations, of which 26 (86.7%) are epistemic and 4 (13.3%) are speech act. In the conversation task, there are 319 non-content relations, of which 32 (8.2%) are epistemic, and 358 (91.8%) are speech act ($\chi^2(1) = 483.84$, $p < .001$).

RELATION BETWEEN SOURCE OF COHERENCE AND ORDER OF THE SEGMENTS. As expected, content relations showed a preference for basic order (87%; binomial test: $p < .05$) and non-content relations showed a preference for non-basic order (78%; binomial test: $p < .05$).

2.3. Discussion

We did not find evidence for the hypothesis that causal, negative and negative causal relations were more difficult for younger children. Does this imply that complex relations are not acquired later, contrary to what was expected on the basis of the classification? We believe there is another explanation, at least for the use of the causal and negative additive relations: It is likely that even the youngest age group has already acquired causal and negative relations. Evers-Vermeul (2005)

and Evers-Vermeul and Sanders (submitted for publication) have looked at the first occurrence of connectives in naturalistic data from children of age 1;6-up. Although they studied connectives and not coherence relations, their data suggest that causal and contrastive connectives occur very early indeed.[5] Before they are three, most children have acquired causal (*want, omdat*) and negative (*maar*) connectives. Hence, these results would explain why we do not find a difference between the age groups in this study. At the same time, however, these data of younger children show direct evidence for the idea of cumulative cognitive complexity: negative connectives emerge later than positive connectives and causal connectives are preceded by additive connectives.

In the case of negative causal relations, we seem to be dealing with a bottom effect: only 2% of the relations produced was negative causal. Consequently, it is not possible to say whether the relation as such is too complex for even the oldest children or whether the tasks employed did not elicit the use of negative causal relations. This point will be taken up in the next section.

For Source of Coherence, we did find that content relations were indeed simpler than non-content relations. Our data accord in this respect with the suggestions by Bates (1976) and Bloom and Capatides (1987) that content relations are simpler than non-content ones. Yet, they contrast with the findings of Kyratzis et al. (1990), suggesting that their findings (speech act relations are acquired before content relations) were indeed biased by the communicative context of the task they used. This interpretation is strongly corroborated by another new finding in this study: the type of relations children use depends on the communicative task. In a descriptive task they mainly use content relations, while non-content relations predominate in conversational tasks. Finally, our data provide convincing evidence for the status of Order of the Segments as an independent primitive in our classification: non-content relations tend to be realized in a non-basic order, a bias which cannot be accounted for in terms of a general Iconicity principle: a temporal ordering of the events described by the segments is irrelevant in the case of non-content relations.

In conclusion, our data provide only moderate support for the strong prediction. There is no indication that relations that are classified as complex by Sanders et al. (1992) are acquired later. This holds both for causal relations as opposed to additive relations, for negative relations as opposed to positive relations (although here the difference is marginally significant), and for negative causal relations. It seems that even the youngest age group has acquired causal relations, so that differences between age groups do not occur. About half of the relations produced by both age groups are causal. This suggests that in future research younger children should be tested.

3. Study 2: children's use of contrastive cause–consequence relations

The data of Study 1 were inconclusive with respect to the question of complex relations, more specifically, with respect to the children's proficiency in dealing with negative causal relations. Example (13) is a classic example of such a relation: in the terminology of Sanders et al. (1992), this example is an instance of the *Contrastive Cause–Consequence* (CCC) relation (also known as the Denial of Expectation relation).

(13)   Although John is a Republican, he is honest.

According to most analyses of such examples (e.g., Lakoff, 1971; Van Dijk, 1977; Abraham, 1979; Blakemore, 1989; Spooren, 1989; Oversteegen, 1997), a *Contrastive Cause–Consequence*

---

[5] As we have stated repeatedly, this does not necessarily mean that causal coherence relations also occur very early, but that inference is very likely.

relation presents two states of affairs (John being a Republican and John being honest), in which the second state-of-affairs runs counter to what could be expected on the basis of the first. In other words, a *Contrastive Cause–Consequence* relation like (13) can only be used if there is some sort of presupposition or expectation that ''generally Republicans are dishonest''.

Because of the implicational character of the presupposition underlying (13) this relation is classified as causal, and because the implication can only be related to the actual utterance through a negation, the relation is classified as negative. Hence, the *Contrastive Cause-Consequence* relation is an example of a negative causal relation.

Ever since Piaget, it is known that CCC relations, or, as he calls them, relations of discordance, are notoriously difficult to learn. Yet, the precise cause of the complexity is difficult to assess, because the factors that may determine the complexity have not been isolated carefully. For instance, (13) may be complex for syntactic reasons (it is a subordinate construction), for lexical reasons (because of the conjunction *although*), or for semantics reasons (because of the presence of the negation), or for reasons of discourse coherence (the complexity of the relation-itself).

Furthermore, different researchers give different assessments of the exact age at which children have fully acquired the relation. Some put it as early as age 7 (Bloom et al., 1980; Peterson, 1986), others put it as late as age 13 (Piaget, 1924/1969; Katz and Brent, 1968; Scott, 1984). In spontaneous observational data, the CCC relation is hardly ever found. This is partly due to the fact that most researchers focus on connectives rather than coherence relations, and the connective *although* is very infrequent in child language. But also in the research discussed above, explicitly aimed at analyzing relations rather than connectives, there were few occurrences of a CCC relation. This makes it uncertain whether children do not have the proficiency to use the relation or whether the communicative context in which the corpus was elicited did not give rise to the use CCC relations. Because of the infrequency of CCC relations in spontaneous speech, the question of the child's proficiency in dealing with these relations needed to be answered via an experiment.

The experiment was designed to investigate the children's proficiency with respect to the CCC relation.[6] What we wanted to know was whether children display developmental patterns, and if so, whether these patterns reflect the analysis of coherence relations in terms of the categorization. In the experiment, children had to complete forced-choice items. In those items, they had to choose between two possible completions of a discourse fragment. The items were designed in such a manner that the choice reflects understanding of the coherence relation.

The question under study is whether CCC relations are more complex than other relations. This leads to two hypotheses:

Hypothesis 1. Children perform worse on items involving CCC relations than on items involving non-causal and/or non-negative relations.

Hypothesis 2. Younger children perform worse on items involving CCC relations than older children.

3.1. Method

3.1.1. Material
Two sorts of item sets were developed. In the design of the sets, the analysis of CCC relations implying (presupposing) a general rule was made use of. The first set consisted of items like the following.

---

[6] The experiment was designed in close collaboration with and carried out by Hanny den Ouden (cf. Den Ouden, 1995).

(14)   Contrastive Consequence–Cause (negative causal: sentence completion)
        People who are lok wear sweaters.
        Els is lok but she is . . .
        a. not wearing a sweater.
        b. wearing a sweater.

That is, the subjects were given a general rule, and on the basis of this rule they had to choose between two solutions to complete a sequence containing *but*. This task was chosen because it more or less resembles the production process: On the basis of an available rule, the speaker makes a formulation choice. The items contained nonsense words (e.g., the word *lok*) to prevent the children to base their choice on real world knowledge in stead of on knowledge of the linguistic connective *but* and the coherence relation it encodes.

The second set of items was of the following type:

(15)   Contrastive Consequence–Cause (negative causal: inferring a general rule)
        Manel has a lot of money, but she is not bap.
        a. People who have much money are not bap.
        b. People who have much money are bap.

In this case the subjects were given a *but*-sequence. On the basis of this sequence, they had to choose which of two alternatives they thought to be the case as a rule. In a way, this task resembles the interpretation process: The underlying idea is that a correct interpretation of a CCC relation implies inferring the general rule.

Hypothesis 1 predicts that CCC relations are more complex than positive causal relations and negative additive relations. To test this prediction, other items were used that contain *Result* relations with the marker *and therefore* and *Contrast* relations with the marker *but*. Examples are:

(16)   Result relation (positive causal): sentence completion
        Tall people are rast.
        Broes is tall and therefore she is . . .
        a. rast
        b. not rast
(17)   Result relation (positive causal): inferring a general rule
        Jani has black hair and therefore he is iepen.
        a. People with black hair are not iepen.
        b. People with black hair are iepen.
(18)   Contrast relation (negative additive): sentence completion
        Jakki goes to ploef, but Boli goes to . . .
        a. pluik
        b. ploef

Because the interpretation of the *Contrast* relation does not imply inferring a general rule (see Spooren, 1989), there were no *Contrast* items in the inference task.

The presence (present/absent) and position of a negation in the items (first clause, second clause), and the position of the nonsense word (first clause/second clause), were systematically varied. In total, there were 24 items in the completion task, and 16 items in the inference task.

Apart from these 40 nonsense items, there were 20 meaningful items, without a nonsense word. These items were added to study whether the children were able to perform the task if they could use real world knowledge. Twelve meaningful items were presented in the completion task, and eight items in the inference task.

3.1.2. Participants

On the basis of a number of pretests, it was decided to use children from grade 3 as the younger age group (younger children did not succeed in performing the task) and children from grade 6 as the older age group. In the experiment, 32 children from grade 3 and 29 children from grade 6 cooperated. The children were recruited from three Dutch elementary schools. The reading level of the children varied equally between age groups. The younger group varied in age from 8;0 to 9;9 years (mean: 8;7), the older group varied in age from 10;11 to 13;0 years (mean: 11;8). In both age groups, the number of boys was somewhat larger than the number of girls (grade 3: 17 boys, grade 6: 16 boys).

3.1.3. Design

The items were presented to the subjects in two blocks. In the first block the nonsense items were presented, in the second block the regular items. This order was fixed to prevent the subjects from developing a strategy of using real world knowledge (as induced by the meaningful items) to respond to the nonsense items.

Each block consisted of two subblocks: a completion task block and an inference task block. The order of these subblocks varied between lists. The order of the items within each subblock was randomized, but did not vary between subjects. The factor Relation Type (for completion task: CCC, Result, Contrast; for inference task: CCC, Result) was repeated with respect to both subjects and items. The factor Age Group (grade 3, grade 6) was a between-subjects, within-items factor.

3.1.4. Procedure and data analysis

The experiment was carried out individually. It consisted of 60 two-choice items. The experiment was preceded by a lengthy instruction to the task. Then the children were presented with test items until it was clear that they understood the task. After that, the 60 experimental items were presented to the children. Presentation of the items differed for the two age groups. For the younger group the test items and the experimental items were presented both orally and in writing, because a pretest had shown that this age group could not cope with a written-only task. The older group were given the test items orally and in writing, but the experimental items only in writing.

In the case of sentence completion items the children chose between two possible continuations, as in example (14). In the case of general inference items the children chose between two possible rules, see example (15).

A percentage correct responses was calculated per type of relation per child. This was done separately for nonsense and meaningful items and for the completion task and the inference task. These percentages were analyzed using Mann–Whitney tests to test differences between age groups and Wilcoxon matched-pairs tests to test differences within age groups.

3.2. Results

The results are summarized in Table 2.

These data will be discussed separately for the nonsense items and for the meaningful items.

Table 2
Percentages correct responses as a function of age, task, and relation type (Study 2)

Task	Completion			Inference	
	NegAdd	PosCau	NegCau	PosCau	NegCau
Items containing nonsense words					
Age					
8–9	79	82	41	75	25
11–12	93	92	62	93	53
Meaningful items					
Age					
8–9	88	95	44	95	59
11–12	97	97	78	100	78

3.2.1. Nonsense items

Hypothesis 1 states that CCC relations are more difficult than the other relations. This was found in both the inference task and the completion task. For the inference task: CCC relations: 39% correct, positive causal relations: 84% correct ($z = 5.56$, $p < .001$; for the completion task: CCC relations: 51% correct, positive causal relations: 86% correct (difference with CCC-relations: $z = 5.10$, $p < .001$), negative additive relations 86% correct (difference with CCC-relations: $z = 5.53$, $p < .001$).

The second hypothesis states that younger children have more trouble with the CCC relations than older children. This was also found for both the inference task (grade 3: 25% correct, grade 6: 53% correct, $z = 2.66$, $p < .01$) and the completion task (grade 3: 41% correct, grade 6: 62% correct, $z = 2.32$, $p < .05$).

3.2.2. Meaningful items

As in the case of the nonsense items, CCC relations were more difficult than other relations, both in the inference task and the completion task (inference task: CCC relations: 68% correct, positive causal relations: 97% correct, $z = 5.30$, $p < .001$; completion task: CCC relations: 60% correct, positive causal relations: 96% correct (difference with CCC relations: $z = 4.95$, $p < .001$, negative additive items: 92% correct (difference with CCC relations: $z = 4.71$, $p < .001$).

It is also the case that younger children find the CCC relation more difficult than the older children (inference task: grade 3: 59% correct, grade 6: 78% correct, $z = 2.37$, $p < .05$; completion task: grade 3: 44% correct, grade 6: 78% correct, $z = 3.45$, $p < .01$).

3.3. Discussion

The results show that both age groups can deal with the concept of polarity and the concept of causality *per se*. Even the younger of the two age groups performs substantially better than chance at both negative additive items and positive causal items. The complexity is in the combination of the two. Items that are both negative AND causal create difficulties. Note that this effect cannot be attributed to the complexity of the marker of the CCC relation because in Contrast items and in CCC items the same marker has been used (namely the marker *but*). Nor can it be attributed to the syntactic or semantic complexity of the constructions, because in all items the same syntactic construction was used, namely a coordinate conjunction, and in all items the presence or absence of a negation was varied systematically. The conclusion must be that it is the complexity of the coherence relation that creates the complexity. And this is precisely in line with the predictions following from our classification of coherence relations.

The same conclusion, namely that negative causal relations are more complex than the other relations in the experiment, can be drawn from the fact that, with age, children learn to interpret the CCC relation better. This may not be so obvious from the mean scores, because these are all relatively low, specifically for the nonsense items. But it is certainly obvious if we look at the individual response patterns of the subjects. For instance there were 13 subjects who gave correct answers to all or nearly all items. Of these, 12 were from the older group and 1 from the younger group. By contrast, the 14 children that seemed to answer randomly nearly all came from the younger age group.

In conclusion, it seems that these data lend support to the thesis that CCC relations are more complex than the others. This is in line with our cognitive theory of coherence relations which analyzes CCC relations as most complex.

4. General discussion

The child language literature provides ample support for both the weak and the strong prediction that follows from the classification of coherence relations as proposed by Sanders et al. (1992). Additive relations are acquired before causal relations and positive relations are acquired before negative relations. New empirical data were gathered in the two studies reported in this paper. Study 1 was an analysis of spoken data in two communicative settings (a descriptive task and a conversational task). It has shown a strong interaction between Source of Coherence and Order of the Segments, thus corroborating these basic concepts. The experiment in Study 2 has shown that negative causal relations are acquired later than negative additive and positive causal relations, irrespective of syntactic construction and semantic or lexical complexity, in line with the strong prediction following from the classification. Therefore, these data support a cognitive interpretation of the classification: Language users make use of their (tacit) knowledge of these basic concepts when they process language.

Do the results of Study 1 really support the idea of cognitive complexity? After all, older children do *not* use complex relations like causals more often than younger children. We believe that this result is not in conflict with the complexity idea, because even the younger children in our study had already acquired the concept of causality. The implication is that the utterances of younger children should be studied to find the kind of differences under discussion. This is corroborated by Evers-Vermeul's (2005) analysis of causal connectives produced by very young Dutch children (ages between 1;6 and 6;0), who predominantly showed the acquisition order additive–temporal–causal.

The ecological validity of the task in Study is admittedly relatively low. The children had to make inferences on the basis of nonsense words. However, we believe that in combination with the more naturalistic tasks in Study 1, this somewhat artificial nature of the task is compensated for. Moreover, the design of Study 2 allows a more precise investigation of questions that remained unanswered when using natural settings: It taps communicative behavior that does not manifest itself in more or less spontaneous settings. The converging nature of these different sets of data allow for the conclusions that we drew.

Nevertheless, sufficient questions for future research remain. Additive relations (or at least the content additive relations) can be further subdivided into purely additive and additive-temporal relations. Studying children's narratives, Berman (1996) found that younger children (3-year olds) use Hebrew *ve (*roughly equivalent to English *and*) as a clause-chaining device, but that older children (5- and 9-year olds) increasingly use it to express temporal sequentiality. They also use relatively more other lexical markers of sequentiality (*then*, *afterwards*, *after that*). This tendency may be expected because temporal sequence is taken to be more complex than purely

additive relations: interpretation involves the temporality feature. Obviously, there are many degrees of temporality (overlap, (partial) precedence, (partial) 'postcedence'), and there are many (mostly formal semantically inspired) studies of how natural language is sensitive to expressing these distinctions (see Oversteegen, 1989 for an analysis of Dutch temporal connectives). The order in which such distinctions are acquired may shed light on the validity of the Sanders et al. (1992) claim that the temporal distinction is secondary to the additive–causal distinction.

As a methodological point, we see the fact that the analysis of discourse coherence crucially involves conceptual relations rather than linguistic connections as an urgent issue for future research. In his paper on the methodology of cognitive linguistics, Langacker (1999) takes as a starting point for linguistic research the analysis of *explicit* linguistic elements. This seems to put the analysis of coherence relations outside of the realm of linguistics, by definition. We believe that such a move would be precipitate and that linguistic regularities behind coherence relations can and should be studied within linguistics. But we are also convinced that the analysis of discourse coherence is in urgent need of reliable and valid analyses of coherence relations. As we have argued elsewhere (Sanders and Spooren, 2007), the strategy of 'converging evidence' followed in this article may be beneficial to that goal.

Acknowledgements

Appendix A. Overview of the classification (Sanders et al., 1993)

Overview of the taxonomy and some prototypical RST-relations

Basic Operation	Source of Coherence	Order	Polarity	Class	Relation in Rhetorical Structure Theory (RST)
Causal	Content	Basic	Positive	1.a	(Non-)volitional cause
				1.b	Condition
Causal	Content	Basic	Negative	2.	Contrast
Causal	Content	Non-basic	Positive	3.a	(Non-)volitional result
				3.b	Condition
Causal	Content	Non-basic	Negative	4.	Contrast
Causal	Epist/SA	Basic	Positive	5.a	Evaluation
				5.b	Interpretation
Causal	Epist/SA	Basic	Negative	6.	Anti-thesis
Causal	Epist/SA	Non-basic	Positive	7.a	Evidence
				7.b	Motivation
Causal	Epist/SA	Non-basic	Negative	8.	Anti-thesis
Additive	Content	–	Positive	9.a	Sequence
				9.b	Joint
Additive	Content	–	Negative	10.	Contrast
Additive	Epist/SA	–	Positive	11.	Presentational sequence
Additive	Epist/SA	–	Negative	12.	Concession

Note. Epist/SA: Epistemic / Speech Act. In the original classifications we distinguished between Semantic and Pragmatic relations, where the latter category included both Epistemic and Speech Act relations. For reasons of presentation (consistency with other literature) we have avoided the original terminology in this article.

Below, each class of relations is illustrated with example of a typical RST-relation. Linguistic markers are in italics.

(1) <u>Non-volitional result</u>
     Because there was a heavy storm, the tiles fell from of the roof.
(2) <u>Contrast</u>
     Although there was a storm, no tiles have fallen from the roof.
(3) <u>Non-volitional cause</u>
     The tiles fell from the roof, *because* there was a heavy storm. ees.
(4) <u>Contrast</u>
     No tiles have fallen from the roof, even though there was a heavy storm.
(5) <u>Interpretation</u>
     A lot of tiles have fallen from the roof. There must have a been a heavy storm.
(6) <u>Anti-thesis</u>
     Although there was a heavy storm, I am sure not a tile will have fallen from the roof.
(7) <u>Evidence</u>
     I am absolute sure there was a heavy storm. All the tiles have fallen from the roof.
(8) <u>Anti-thesis</u>
     I am sure not a tile will have fallen from the roof, *even though* there was a heavy storm.
(9) <u>Sequence</u>
     There was a heavy storm last Tuesday. On wednesday we had snow.
(10) <u>Contrast</u>
     There was a heavy storm last Tuesday. But the next day we had beautiful sunny weather.
(11) <u>Presentational Sequence</u>
     (There was a heavy storm last Tuesday). The tiles have fallen from the roof. *Furthermore*,
     I saw a tree lying on the street.
(12) <u>Concession</u>
     Some tiles have fallen from the roof. *But* I still think the storm was not that bad.

Appendix B. Sample from the corpus (Study 1)

Utterance	Classification				RST-relation
Grade 1					
(a) Die meneer schiet een pijl. En dan klapt de band.	ca	con	ba	po	CA
'That man shoots an arrow. And then the tire bursts.'					
(b) Die mensen huilen allemaal. Omdat die vioolspeler zo mooi speelt.	ca	con	nba	po	RE
'Those people are all crying. Because that violinist is playing so beautifully.'					
(c) Ze zijn ramen aan het wassen. En dan laat hij zijn emmertje vallen.	ad	con	na	po	CI
'They are cleaning the windows. And then he drops his little bucket.'					
(d) We verliezen soms [met voetbal]. Maar meestal winnen we.	ad	con	na	neg	OP
'Sometimes we lose [at soccer]. But usually we win.'					
(e) Ik heb liever een juf. Want een juf is veel liever als een meester.	ca	sa	nba	pos	EV
'I prefer a female teacher. Because a female teacher is much sweeter than a male teacher.'					
(f) Bassie en Adriaan is leuk. Skippy vind ik niet leuk.	ad	sa	na	neg	OP
'Bassie and Adrian [a children's TV program] is nice. Skippy isn't nice.'					

Appendix B (*Continued*)

Utterance	Classification				RST-relation
Grade 6					
(g) Een klein jongetje duwt de vazen van de tafel af. En daardoor vallen de vazen kapot. 'A little boy pushes the vases of the table. And because of that the vases fall to pieces.'	ca	con	ba	pos	CA
(h) Hier denken de mensen dat er een haai komt. Want er zijn twee jochies die houden een stok met een vin omhoog. 'Here the people think that a shark is coming. Because there are two kids who hold up a stick with a fin.'	ca	con	nba	pos	RE
(i) Hier heb je een skelet van een dinosaurus. En daar staan allemaal holbewoners bij. 'Here you have the skeleton of a dinosaure. And a lot of cavemen are standing close to it.'	ad	con	na	pos	CI
(j) Die ene visser heeft een grote boot met allemaal visnetten. En die vangt helemaal geen vis. 'That one fisherman has a large boat with all sorts of fishing nets. And that one isn't catching any fish.'	ca	con	ba	neg	CO
(k) Mij maakt het nooit uit [of ik merkkleding draag]. Als je maar leuke dingen hebt. 'To me it never matters [whether I'm wearing brand clothing. As long as you [i.e. ''I''] have nice things.'	ca	sa	nba	pos	CD
(l) Ik vind het ook wel spannend [om van school te gaan]. Want daar is alles nieuw. 'I think it is also quite exciting [to leave school]. Because everything is new there.'	ca	ep	nba	pos	EV

Note. Classification categories: ad: additive; ca: causal; ba: basic order; nba: non-basic order; na: not applicable; con: content; sa: speech act; ep: epistemic; pos: positive; neg: negative.

Relations: CA: cause; CI: circumstance; OP: opposition; RE: result; CO: contrast; CD: condition; EV: evidence.

Appendix C. Analysis of content, epistemic and speech act relations (Study 1)

The paraphrase criteria sketched below were introduced in Sanders (1997). A relation was considered a content relation if it passed the following paraphrase test:

(C1)   Paraphrase criteria for content relations
      a.   S1. This fact causes/caused the fact S2.
      b.   S1. This fact is/was caused by the fact that S2.

The categorization as epistemic or speech act relations was done with a satisfactory amount of interrater agreement (Cohen's $\kappa = .69$, $n = 430$). Non-content relations were considered epistemic if they could be paraphrased with any of the following phrases:

(C2)   Paraphrase criteria for epistemic relations
      a.   S1. And on the basis of this I conclude here and now that S2.
      b.   I conclude here and now that S1. I do this on the basis of S2.
      c.   It must be the case that S1. My evidence for that is S2.
      d.   It probably is the case that S1. My evidence for that is S2.

Another distinguishing feature of epistemic relations is that they involve the generation/inference of *new* knowledge (as opposed to the conveying of existing knowledge that occurs frequently in speech act relations). A typical epistemic relation is:

(C3)   [Chantal points to a mummy walking around; in the corner of the picture there is an open coffin]
   a.   Daar komt een mummie
       'There comes a mummy'
   b.   Misschien hebben ze de kist opengedaan
       'Maybe they opened the coffin'

(Chantal 12:6)

The relationship between (C2)a and (C2)b is epistemic because the speaker concludes here and now that a possible explanation for the mummy walking around is that someone opened the coffin.

Relationships were categorized as speech act relations if they could be paraphrased as in (C3).

(C3)   Paraphrase criteria for speech act relations
   a.   I say to you, order you, ask you ... (etc.) S1. The basis for my speech act is that S2.
   b.   Given that S1 is the case, I say to you, order you, ask you ... (etc.) S2.

A typical example of a speech act relation is the following:

(C4)   a.   En Sesamstraat vind ik niet meer zo leuk
       'And I don't like Sesame Street very much anymore'
   b.   Want daar ben ik nu te groot voor.
       'Because I'm too old for that'

(Rose 12:3)

The relation between (C4)a. and b. is a speech act relation because the speaker is telling the interviewer that she does not like Sesame Street very much anymore, and the basis for this statement is that she has outgrown the program.

References

Abraham, Werner, 1979. But. Studia Linguistica 33, 89–119.

Asher, Nicholas, Lascarides, Alex, 1998. Bridging. Journal of Semantics 15 (1), 83–113.

Bates, Elisabeth A., 1976. Language in Context: The Acquisition of Pragmatics. Academic Press, New York.

Bebout, Linda J., Segalowitz, Sidney J., White, G.J., 1980. Children's comprehension of causal constructions with BECAUSE and SO. Child Development 51, 565–568.

Berman, Ruth A., 1996. Form and function in developing narrative abilities: the case of AND. In: Slobin, D., Gerhardt, D.J., Kyratzis, A., Guo, J. (Eds.), Social Interaction, Context and Language. Erlbaum, Mahwah, NJ, pp. 243–268.

Blakemore, Diana, 1989. Denial and contrast: a Relevance Theoretic analysis of BUT. Linguistics and Philosophy 12, 15–37.

Bloom, Lois, Capatides, Joanne B., 1987. Sources of meaning in the acquisition of complex syntax: the sample of causality. Journal of Experimental Child Psychology 43, 112–128.

Bloom, Lois, Lahey, Margaret, Hood, Lois, Lifter, Karin, Fiess, Kathleen, 1980. Complex sentences: acquisition of syntactic connectives and the semantic relations they encode. Journal of Child Language 7, 235–261.

Clark, Herb H., Clark, Eve V., 1977. Psychology and Language: an Introduction to Psycholinguistics. Harcourt Brace Jovanovich, New York.

Corrigan, Roberta L., 1975. A scalogram analysis of the development of the use and comprehension of BECAUSE in children. Child Development 46, 195–201.

Dancygier, Barbara, Sweetser, Eve E., 2000. Constructions with IF, SINCE and BECAUSE: Causality, epistemic stance, and clause order. In: Couper-Kuhlen, E., Kortmann, B. (Eds.), CAUSE–CONDITION–CONCESSION–CONTRAST. Cognitive and Discourse Perspectives. Mouton de Gruyter, Berlin, pp. 111–142.

Den Ouden, Hanny, 1995. De verwerving van de negatief causale relatie [The acquisition of the negative causal relation]. Toegepaste Taalwetenschap in Artikelen 53, 83–94.

Eisenberg, Ann R., 1980. A syntactic, semantic and pragmatic analysis of conjunction. Stanford Papers and Reports on Child Language Development 19, 70–78.

Evers-Vermeul, Jacqueline, 2005. Connections Between Form and Function of Dutch Connectives. Change and Acquisition as Windows on Form-Function Relations. PhD Thesis, UiL OTS, Utrecht University.

Evers-Vermeul, Jacqueline, Sanders, Ted. The emergence of Dutch connectives; how cumulative cognitive complexity explains the order of acquisition, submitted for publication.

French, Lucia A., Nelson, Katherine, 1985. Young Children's Knowledge of Relational Terms: Some Ifs, Ors and Buts. Springer Verlag, New York.

Garnham, Alan, Oakhill, Jane (Eds.), 1992. Discourse Representation and Text Processing. A Special Issue of Language and Cognitive Processes. Lawrence Erlbaum Associates, Hove, UK.

Gernsbacher, Morton A., Givón, Talmy (Eds.), 1995. Coherence in Spontaneous Text. Benjamins, Amsterdam.

Graesser, Arthur C., McNamara, Danielle S., Louwerse, Max M., 2003. What do readers need to learn in order to process coherence relations in narrative and expository text? In: Sweet, A.P., Snow, C.E. (Eds.), Rethinking Reading Comprehension. Guilford, New York, pp. 82–98.

Graesser, Arthur C., Millis, Keith K., Zwaan, Rolf, 1997. Discourse comprehension. Annual Review of Psychology 48, 163–189.

Halliday, Michal A.K., Hasan, Ruqaiya, 1976. Cohesion in English. Longman, London.

Handford, Martin, 1991. Waar is Wally? [Where's Waldo?] M&P Uitgeverij bv., Weert.

Hickmann, Maya, 2003. Children's Discourse: Person, Space and Time Across Languages. Cambridge University Press, Cambridge, UK.

Hobbs, Jerry R., 1979. Coherence and coreference. Cognitive Science 3, 67–90.

Hood, Lois, Bloom, Lois, 1979. What, When, and How about Why: A Longitudinal Study of Early Expressions of Causality. Monographs of the Society for Research in Child Development 44 (Serial No. 181).

Katz, Evelyn W., Brent, Sandor D., 1968. Understanding connectives. Journal of Verbal Learning and Verbal Behavior 7, 501–509.

Kyratzis, Amy, Guo, Jiansheng, Ervin-Tripp, Susan, 1990. Pragmatic conventions influencing children's use of causal constructions in natural discourse. In: Hall, K., Koenig, J.P., Meacham, M., Reinman, S., Sutton, L.A. (Eds.), Proceedings of the 16th Annual Meeting of the Berkeley Linguistics Society, BLS, Berkeley, CA, pp. 205–214.

Lakoff, Robin, 1971. If's, and's, and but's about conjunction. In: Fillmore, C.J., Langendoen, D.T. (Eds.), Studies in Linguistic Semantics. Holt, Rhinehart & Winston, New York, pp. 114–149.

Langacker, Ronald, 1999. Assessing the cognitive linguistic enterprise. In: Janssen Th., A.J.M., Redeker, G. (Eds.), Cognitive Linguistics: Foundations, Scope and Methodology. Mouton de Gruyter, Berlin, Cognitive Linguistics Research 15., pp. 13–59.

Mann, William C., Thompson, Sandra A., 1988. Rhetorical Structure Theory: toward a functional theory of text organization. Text 8, 243–281.

McCabe, Allyssa, Peterson, Carole, 1985. A naturalistic study of the production of causal connectives by children. Journal of Child Language 12, 145–159.

Oversteegen, Eleonore, 1989. Tracking Time, a Proposal for the Representation of Temporal Expressions in Dutch. PhD Thesis, University of Utrecht.

Oversteegen, Eleonore, 1997. On the pragmatic nature of causal and contrastive concessives. Discourse Processes 24, 51–86.

Peterson, Carole, 1986. Semantic and pragmatic uses of BUT. Journal of Child Language 13, 583–590.

Peterson, Carole, McCabe, Allyssa, 1987. The connective AND: do older children use it less as they learn connectives? Journal of Child Language 14, 375–381.

Piaget, Jean, 1969. Judgement and Reasoning in the Child. Routledge & Kegan Paul, London. [Translation of: Jean Piaget, 1924. Le jugement et le raisonnement chez l'enfant. Delachaux et Niestlé, Neuchatel, etc.].

Piaget, Jean, 1971. The Language and Thought of the Child. Routledge & Kegan Paul, London. [Translation of: Jean Piaget, 1923. Le langage et le pensée chez l'enfant. Delachaux et Niestlé, Neuchatel, etc.].

Sanders, Ted, 1997. Semantic and pragmatic sources of coherence. On the categorization of coherence relations in context. Discourse Processes 24, 119–147.

Sanders, Ted, Pander Maat, Henk, 2006. Cohesion and coherence: Linguistic approaches. In: Brown, K., et al. (Eds.), Encyclopedia of Language and Linguistics. 2nd ed. Elsevier, London.

Sanders, Ted, Spooren, Wilbert, 2001. Text representation as an interface between language and its users. In: Sanders, T., Schilperoord, J., Spooren, J. (Eds.), Text Representation: Linguistic and Psycholinguistic Aspects. Benjamins, Amsterdam, pp. 1–25.

Sanders, Ted, Spooren, Wilbert, 2007. Discourse and text structure. In: Geeraerts, D., Cuykens, H. (Eds.), Handbook of Cognitive Linguistics. Oxford University Press, Oxford, pp. 916–943.

Sanders, Ted, Spooren, Wilbert, Noordman, Leo, 1992. Toward a taxonomy of coherence relations. Discourse Processes 15, 1–35.

Sanders, Ted, Spooren, Wilbert, Noordman, Leo, 1993. Coherence relations in a cognitive theory of discourse representation. Cognitive Linguistics 4, 93–134.

Scott, Cheryl M., 1984. Adverbial connectivity in conversations of children 6 to 12. Journal of Child Language 11, 423–452.

Schiffrin, Deborah, 1986. Functions of and in discourse. Journal of Pragmatics 10, 41–66.

Spooren, Wilbert, 1989. Some Aspects of the Form and Interpretation of Global Contrastive Coherence Relations. Unpublished Ph.D. Dissertation, K.U. Nijmegen.

Sweetser, Eve E., 1990. From Etymology to Pragmatics. Cambridge University Press, Cambridge.

Taboada, Maite, Mann, William, 2006. Applications of Rhetorical Structure Theory. Discourse Studies 8 (4), 567–588.

Tabor, Whitney, Traugott, Elizabeth C., 1998. Structural scope expansion and grammaticalization. In: Ramat, A.G., Hopper, P.J. (Eds.), The Limits of Grammaticalization. Benjamins, Amsterdam, pp. 227–270.

Traugott, Elizabeth C., 1997. Review of Christian Lehmann, thoughts on grammaticalization. Journal of Linguistics 33, 171–211.

Van Dijk, Teun A., 1977. Connectives in text grammar and text logic. In: Van Dijk, T.A., Petöfi, J. (Eds.), Grammars and Descriptions: Studies in Text Theory and Text Analysis. De Gruyter, Berlin, pp. 11–63.

Verhagen, Arie, 2000. Achter het Nederlands. Over interacties tussen taal en achtergrondcognitie [Beyond Dutch: on interactions between language and background cognition]. Inaugural Speech, Leiden University.

Wing, Clara S., Scholnick, Ellin K., 1981. Children's comprehension of pragmatic concepts expressed in BECAUSE, ALTHOUGH, IF and UNLESS. Journal of Child Language 8, 347–365.

Wilbert Spooren (born 1956, Ph.D. Nijmegen University, 1989) is professor of Language and Communication at the Faculty of Arts of Vrije Universiteit Amsterdam, The Netherlands. His group teaches language and communication at several BA- and MA-levels. His research focuses on the relation between genre, discourse structure and text quality. He has published on these issues nationally and internationally in journals and edited volumes, see www.let.vu.nl/staf/w.spooren.

Ted Sanders (born 1963, Ph.D. Tilburg University, 1992) is professor of Dutch language use and discourse studies at Utrecht University, The Netherlands. His group teaches language and communication at several BA- and MA-levels. His research focuses on discourse structure and coherence, from a linguistic and a psycholinguistic perspective. He is the (co-)author of many journal publications and the (co-)editor of several edited books and special issues, see http://www.let.uu.nl/~ted.sanders/personal/index.php.